



SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Soreq, Hermona
Zakut, Haim
Shani, Moshe

(ii) TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
ANTICHOLINESTERASE SUBSTANCES

(iii) NUMBER OF SEQUENCES: 27

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Kohn & Associates
(B) STREET: 30500 Northwestern Hwy, Suite 410
(C) CITY: Farmington Hills
(D) STATE: Michigan
(E) COUNTRY: US
(F) ZIP: 48334

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: PCT/US95/02806
(B) FILING DATE: 28 February 1995
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Kohn, Kenneth I.
(B) REGISTRATION NUMBER: 30,955
(C) REFERENCE/DOCKET NUMBER: P-307 (Mulford)

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (810) 539-5050
(B) TELEFAX: (810) 539-5055

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2256 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CCTCTCTCCC CTCATCTTTG CCAACCTGCC CCACCTCCTC TGCAGCTGAG CGATAACCCT
60

TGGGCCGACA GTGCCCTAAT CTCCTCCCTC CTGGCTTCTC GACCGACCCT TCACCCTTTC
120

CCTTTCTTTC TCCCAGCAGA CGCCGCCTGC CCTGCAGCCA TGAGGCCCCC GCAGTGTCTG
180

CTGCACACGC CTTCCCTGGC TTCCCCACTC CTTCTCCTCC TCCTCTGGCT CCTGGGTGGA
240

GGAGTGGGGG CTGAGGGCCG GGAGGATGCA GAGCTGCTGG TGACGGTGCG TGGGGGCCGG
300

CTGCGGGGCA TTCGCCTGAA GACCCCCGGG GGCCCTGTCT CTGCTTTCCT GGGCATCCCC
360

TTTGCGGAGC CACCCATGGG ACCCCGTCGC TTTCTGCCAC CGGAGCCCAA GCAGCCTTGG
420

TCAGGGGTGG TAGACGCTAC AACCTTCCAG AGTGTCTGCT ACCAATATGT GGACACCCTA
480

TACCCAGGTT TTGAGGGCAC CGAGATGTGG AACCCCAACC GTGAGCTGAG CGAGGACTGC
540

CTGTACCTCA ACGTGTGGAC ACCATACCCC CGGCCTACAT CCCCCACCCC TGTCTCGTC
600

TGGATCTATG GGGGTGGCTT CTACAGTGGG GCCTCCTCCT TGGACGTGTA CGATGGCCCG
660

TTCTTGGTAC AGGCCGAGAG GACTGTGCTG GTGTCCATGA ACTACCGGGT GGGAGCCTTT
720

GGCTTCCTGG CCCTGCCGGG GAGCCGAGAG GCCCCGGGCA ATGTGGGTCT CCTGGATCAG
780

AGGCTGGCCC TGCAGTGGGT GCAGGAGAAC GTGGCAGCCT TCGGGGGTGA CCCGACATCA
840

GTGACGCTGT TTGGGGAGAG CGCGGGAGCC GCCTCGGTGG GCATGCACCT GCTGTCCCCG
900

CCCAGCCGGG GCCTGTTCCA CAGGGCCGTG CTGCAGAGCG GTGCCCCCAA TGGACCCTGG
960

GCCACGGTGG GCATGGGAGA GGCCCGTCGC AGGGCCACGC AGCTGGCCCA CCTTGTGGGC
1020

TGTCCTCCAG GCGGCACTGG TGGGAATGAC ACAGAGCTGG TAGCCTGCCT TCGGACACGA
1080

CCAGCGCAGG TCCTGGTGAA CCACGAATGG CACGTGCTGC CTCAAGAAAG CGTCTTCCGG
1140

TTCTCCTTCG TGCCTGTGGT AGATGGAGAC TTCCTCAGTG ACACCCCAGA GGCCCTCATC
1200

AACGCGGGAG ACTTCCACGG CCTGCAGGTG CTGGTGGGTG TGGTGAAGGA TGAGGGCTCG
1260

TATTTTCTGG TTTACGGGGC CCCAGGCTTC AGCAAAGACA ACGAGTCTCT CATCAGCCGG
1320

GCCGAGTTCC TGGCCGGGGT GCGGGTCGGG GTTCCCCAGG TAAGTGACCT GGCAGCCGAG
1380

GCTGTGGTCC TGCATTACAC AGACTGGCTG CATCCCGAGG ACCCGGCACG CCTGAGGGAG
1440

GCCCTGAGCG ATGTGGTGGG CGACCACAAT GTCGTGTGCC CCGTGGCCCA GCTGGCTGGG
1500

CGACTGGCTG CCCAGGGTGC CCGGGTCTAC GCCTACGTCT TTGAACACCG TGCTTCCACG
1560

CTCTCCTGGC CCCTGTGGAT GGGGGTGCCC CACGGCTACG AGATCGAGTT CATCTTTGGG
1620

ATCCCCCTGG ACCCCTCTCG AAATAACACG GCAGAGGAGA AAATCTTCGC CCAGCGACTG
1680

ATGCGATACT GGGCCAACTT TGCCCGCACA GGGGATCCCA ATGAGCCCCG AGACCCCAAG
1740

GCCCCACAAT GGCCCCCGTA CACGGCGGGG GCTCAGCAGT ACGTTAGTCT GGACCTGCGG
1800

CCGCTGGAGG TGCGGCGGGG GCTGCGCGCC CAGGCCTGCG CCTTCTGGAA CCGCTTCCTC
1860

CCCAAATTGC TCAGCGCCAC CGACACGCTC GACGAGGCGG AGCGCCAGTG GAAGGCCGAG
1920

TTCCACCGCT GGAGCTCCTA CATGGTGCAC TGGAAGAACC AGTTCGACCA CTACAGCAAG
1980

CAGGATCGCT GCTCAGACCT GTGACCCCGG CGGGACCCCC ATGTCCTCCG CTCCGCCCGG
2040

CCCCTAGCT GTATATACTA TTTATTTTCTAG GGCTGGGCTA TAACACAGAC GAGCCCCAGA
2100

CTCTGCCCAT CCCCACCCCA CCCCACGTC CCCCAGGGGCT CCCGGTCCTC TGGCATGTCT
2160

TCAGGCTGAG CTCCTCCCCG CGTGCCTTCG CCCTCTGGCT GCAAATAAAC TGTTACAGGC
2220

CAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAA

2256

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 614 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Arg Pro Pro Gln Cys Leu Leu His Thr Pro Ser Leu Ala Ser Pro
1 5 10 15

Leu Leu Leu Leu Leu Leu Trp Leu Leu Gly Gly Gly Val Gly Ala Glu
20 25 30

Gly Arg Glu Asp Ala Glu Leu Leu Val Thr Val Arg Gly Gly Arg Leu
35 40 45

Arg Gly Ile Arg Leu Lys Thr Pro Gly Gly Pro Val Ser Ala Phe Leu
50 55 60

Gly Ile Pro Phe Ala Glu Pro Pro Met Gly Pro Arg Arg Phe Leu Pro
65 70 75 80

Pro Glu Pro Lys Gln Pro Trp Ser Gly Val Val Asp Ala Thr Thr Phe
85 90 95

Gln Ser Val Cys Tyr Gln Tyr Val Asp Thr Leu Tyr Pro Gly Phe Glu
100 105 110

Gly Thr Glu Met Trp Asn Pro Asn Arg Glu Leu Ser Glu Asp Cys Leu
115 120 125

Tyr Leu Asn Val Trp Thr Pro Tyr Pro Arg Pro Thr Ser Pro Thr Pro
130 135 140

Val Leu Val Trp Ile Tyr Gly Gly Gly Phe Tyr Ser Gly Ala Ser Ser
145 150 155 160

Leu Asp Val Tyr Asp Gly Arg Phe Leu Val Gln Ala Glu Arg Thr Val
165 170 175

Leu Val Ser Met Asn Tyr Arg Val Gly Ala Phe Gly Phe Leu Ala Leu
180 185 190

Pro Gly Ser Arg Glu Ala Pro Gly Asn Val Gly Leu Leu Asp Gln Arg
195 200 205

Leu Ala Leu Gln Trp Val Gln Glu Asn Val Ala Ala Phe Gly Gly Asp
210 215 220

Pro Thr Ser Val Thr Leu Phe Gly Glu Ser Ala Gly Ala Ala Ser Val
225 230 235 240

Gly Met His Leu Leu Ser Pro Pro Ser Arg Gly Leu Phe His Arg Ala
245 250 255

Val Leu Gln Ser Gly Ala Pro Asn Gly Pro Trp Ala Thr Val Gly Met
260 265 270

Gly Glu Ala Arg Arg Arg Ala Thr Gln Leu Ala His Leu Val Gly Cys
275 280 285

Pro Pro Gly Gly Thr Gly Gly Asn Asp Thr Glu Leu Val Ala Cys Leu
290 295 300

Arg Thr Arg Pro Ala Gln Val Leu Val Asn His Glu Trp His Val Leu
305 310 315 320

Pro Gln Glu Ser Val Phe Arg Phe Ser Phe Val Pro Val Val Asp Gly
325 330 335

Asp Phe Leu Ser Asp Thr Pro Glu Ala Leu Ile Asn Ala Gly Asp Phe
340 345 350

His Gly Leu Gln Val Leu Val Gly Val Val Lys Asp Glu Gly Ser Tyr
355 360 365

Phe Leu Val Tyr Gly Ala Pro Gly Phe Ser Lys Asp Asn Glu Ser Leu
370 375 380

Ile Ser Arg Ala Glu Phe Leu Ala Gly Val Arg Val Gly Val Pro Gln
385 390 395 400

Val Ser Asp Leu Ala Ala Glu Ala Val Val Leu His Tyr Thr Asp Trp
405 410 415

Leu His Pro Glu Asp Pro Ala Arg Leu Arg Glu Ala Leu Ser Asp Val
420 425 430

Val Gly Asp His Asn Val Val Cys Pro Val Ala Gln Leu Ala Gly Arg
435 440 445

Leu Ala Ala Gln Gly Ala Arg Val Tyr Ala Tyr Val Phe Glu His Arg
450 455 460

Ala Ser Thr Leu Ser Trp Pro Leu Trp Met Gly Val Pro His Gly Tyr
465 470 475 480

Glu Ile Glu Phe Ile Phe Gly Ile Pro Leu Asp Pro Ser Arg Asn Tyr
485 490 495

Thr Ala Glu Glu Lys Ile Phe Ala Gln Arg Leu Met Arg Tyr Trp Ala
500 505 510

Asn Phe Ala Arg Thr Gly Asp Pro Asn Glu Pro Arg Asp Pro Lys Ala
515 520 525

Pro Gln Trp Pro Pro Tyr Thr Ala Gly Ala Gln Gln Tyr Val Ser Leu
530 535 540

Asp Leu Arg Pro Leu Glu Val Arg Arg Gly Leu Arg Ala Gln Ala Cys
545 550 555 560

Ala Phe Trp Asn Arg Phe Leu Pro Lys Leu Leu Ser Ala Thr Asp Thr
565 570 575

Leu Asp Glu Ala Glu Arg Gln Trp Lys Ala Glu Phe His Arg Trp Ser
580 585 590

Ser Tyr Met Val His Trp Lys Asn Gln Phe Asp His Tyr Ser Lys Gln
595 600 605

Asp Arg Cys Ser Asp Leu
610

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3096 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 160..1959

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CCTCTCTCCC CTCATCTTTG CCAACCTGCC CCACCTCCTC TGCAGCTGAG CGATAACCCT
60

TGGGCCGACA GTGCCCTAAT CTCCTCCCTC CTGGCTTCTC GACCGACCCT TCACCCTTTC
120

CCTTTCTTTC TCCCAGCAGA CGCCGCCTGC CCTGCAGCC ATG AGG CCC CCG CAG 174
Met Arg Pro Pro Gln
1 5

TGT CTG CTG CAC ACG CCT TCC CTG GCT TCC CCA CTC CTT CTC CTC CTC 222
Cys Leu Leu His Thr Pro Ser Leu Ala Ser Pro Leu Leu Leu Leu
10 15 20

CTC TGG CTC CTG GGT GGA GGA GTG GGG GCT GAG GGC CGG GAG GAT GCA 270
Leu Trp Leu Leu Gly Gly Gly Val Gly Ala Glu Gly Arg Glu Asp Ala
25 30 35

GAG CTG CTG GTG ACG GTG CGT GGG GGC CGG CTG CGG GGC ATT CGC CTG 318
Glu Leu Leu Val Thr Val Arg Gly Gly Arg Leu Arg Gly Ile Arg Leu
40 45 50

AAG ACC CCC GGG GGC CCT GTC TCT GCT TTC CTG GGC ATC CCC TTT GCG 366
Lys Thr Pro Gly Gly Pro Val Ser Ala Phe Leu Gly Ile Pro Phe Ala
55 60 65

GAG CCA CCC ATG GGA CCC CGT CGC TTT CTG CCA CCG GAG CCC AAG CAG 414
Glu Pro Pro Met Gly Pro Arg Arg Phe Leu Pro Pro Glu Pro Lys Gln
70 75 80 85

CCT TGG TCA GGG GTG GTA GAC GCT ACA ACC TTC CAG AGT GTC TGC TAC 462
Pro Trp Ser Gly Val Val Asp Ala Thr Thr Phe Gln Ser Val Cys Tyr
90 95 100

CAA TAT GTG GAC ACC CTA TAC CCA GGT TTT GAG GGC ACC GAG ATG TGG 510
Gln Tyr Val Asp Thr Leu Tyr Pro Gly Phe Glu Gly Thr Glu Met Trp
105 110 115

AAC CCC AAC CGT GAG CTG AGC GAG GAC TGC CTG TAC CTC AAC GTG TGG 558
Asn Pro Asn Arg Glu Leu Ser Glu Asp Cys Leu Tyr Leu Asn Val Trp
120 125 130

ACA CCA TAC CCC CGG CCT ACA TCC CCC ACC CCT GTC CTC GTC TGG ATC 606
Thr Pro Tyr Pro Arg Pro Thr Ser Pro Thr Pro Val Leu Val Trp Ile
135 140 145

TAT GGG GGT GGC TTC TAC AGT GGG GCC TCC TCC TTG GAC GTG TAC GAT 654
Tyr Gly Gly Gly Phe Tyr Ser Gly Ala Ser Ser Leu Asp Val Tyr Asp
150 155 160 165

GGC CGC TTC TTG GTA CAG GCC GAG AGG ACT GTG CTG GTG TCC ATG AAC 702
Gly Arg Phe Leu Val Gln Ala Glu Arg Thr Val Leu Val Ser Met Asn
170 175 180

TAC CGG GTG GGA GCC TTT GGC TTC CTG GCC CTG CCG GGG AGC CGA GAG 750
Tyr Arg Val Gly Ala Phe Gly Phe Leu Ala Leu Pro Gly Ser Arg Glu
185 190 195

GCC CCG GGC AAT GTG GGT CTC CTG GAT CAG AGG CTG GCC CTG CAG TGG 798
Ala Pro Gly Asn Val Gly Leu Leu Asp Gln Arg Leu Ala Leu Gln Trp

| 200 | 205 | 210 | |
|---|-----|-----|------|
| GTG CAG GAG AAC GTG GCA GCC TTC GGG GGT GAC CCG ACA TCA GTG ACG | | | 846 |
| Val Gln Glu Asn Val Ala Ala Phe Gly Gly Asp Pro Thr Ser Val Thr | | | |
| 215 | 220 | 225 | |
| CTG TTT GGG GAG AGC GCG GGA GCC GCC TCG GTG GGC ATG CAC CTG CTG | | | 894 |
| Leu Phe Gly Glu Ser Ala Gly Ala Ala Ser Val Gly Met His Leu Leu | | | |
| 230 | 235 | 240 | 245 |
| TCC CCG CCC AGC CGG GGC CTG TTC CAC AGG GCC GTG CTG CAG AGC GGT | | | 942 |
| Ser Pro Pro Ser Arg Gly Leu Phe His Arg Ala Val Leu Gln Ser Gly | | | |
| 250 | 255 | 260 | |
| GCC CCC AAT GGA CCC TGG GCC ACG GTG GGC ATG GGA GAG GCC CGT CGC | | | 990 |
| Ala Pro Asn Gly Pro Trp Ala Thr Val Gly Met Gly Glu Ala Arg Arg | | | |
| 265 | 270 | 275 | |
| AGG GCC ACG CAG CTG GCC CAC CTT GTG GGC TGT CCT CCA GGC GGC ACT | | | 1038 |
| Arg Ala Thr Gln Leu Ala His Leu Val Gly Cys Pro Pro Gly Gly Thr | | | |
| 280 | 285 | 290 | |
| GGT GGG AAT GAC ACA GAG CTG GTA GCC TGC CTT CGG ACA CGA CCA GCG | | | 1086 |
| Gly Gly Asn Asp Thr Glu Leu Val Ala Cys Leu Arg Thr Arg Pro Ala | | | |
| 295 | 300 | 305 | |
| CAG GTC CTG GTG AAC CAC GAA TGG CAC GTG CTG CCT CAA GAA AGC GTC | | | 1134 |
| Gln Val Leu Val Asn His Glu Trp His Val Leu Pro Gln Glu Ser Val | | | |
| 310 | 315 | 320 | 325 |
| TTC CGG TTC TCC TTC GTG CCT GTG GTA GAT GGA GAC TTC CTC AGT GAC | | | 1182 |
| Phe Arg Phe Ser Phe Val Pro Val Val Asp Gly Asp Phe Leu Ser Asp | | | |
| 330 | 335 | 340 | |
| ACC CCA GAG GCC CTC ATC AAC GCG GGA GAC TTC CAC GGC CTG CAG GTG | | | 1230 |
| Thr Pro Glu Ala Leu Ile Asn Ala Gly Asp Phe His Gly Leu Gln Val | | | |
| 345 | 350 | 355 | |
| CTG GTG GGT GTG GTG AAG GAT GAG GGC TCG TAT TTT CTG GTT TAC GGG | | | 1278 |
| Leu Val Gly Val Val Lys Asp Glu Gly Ser Tyr Phe Leu Val Tyr Gly | | | |
| 360 | 365 | 370 | |
| GCC CCA GGC TTC AGC AAA GAC AAC GAG TCT CTC ATC AGC CGG GCC GAG | | | 1326 |
| Ala Pro Gly Phe Ser Lys Asp Asn Glu Ser Leu Ile Ser Arg Ala Glu | | | |
| 375 | 380 | 385 | |
| TTC CTG GCC GGG GTG CGG GTC GGG GTT CCC CAG GTA AGT GAC CTG GCA | | | 1374 |
| Phe Leu Ala Gly Val Arg Val Gly Val Pro Gln Val Ser Asp Leu Ala | | | |
| 390 | 395 | 400 | 405 |
| GCC GAG GCT GTG GTC CTG CAT TAC ACA GAC TGG CTG CAT CCC GAG GAC | | | 1422 |
| Ala Glu Ala Val Val Leu His Tyr Thr Asp Trp Leu His Pro Glu Asp | | | |
| 410 | 415 | 420 | |

CCG GCA CGC CTG AGG GAG GCC CTG AGC GAT GTG GTG GGC GAC CAC AAT 1470
 Pro Ala Arg Leu Arg Glu Ala Leu Ser Asp Val Val Gly Asp His Asn
 425 430 435

GTC GTG TGC CCC GTG GCC CAG CTG GCT GGG CGA CTG GCT GCC CAG GGT 1518
 Val Val Cys Pro Val Ala Gln Leu Ala Gly Arg Leu Ala Ala Gln Gly
 440 445 450

GCC CGG GTC TAC GCC TAC GTC TTT GAA CAC CGT GCT TCC ACG CTC TCC 1566
 Ala Arg Val Tyr Ala Tyr Val Phe Glu His Arg Ala Ser Thr Leu Ser
 455 460 465

TGG CCC CTG TGG ATG GGG GTG CCC CAC GGC TAC GAG ATC GAG TTC ATC 1614
 Trp Pro Leu Trp Met Gly Val Pro His Gly Tyr Glu Ile Glu Phe Ile
 470 475 480 485

TTT GGG ATC CCC CTG GAC CCC TCT CGA AAC TAC ACG GCA GAG GAG AAA 1662
 Phe Gly Ile Pro Leu Asp Pro Ser Arg Asn Tyr Thr Ala Glu Glu Lys
 490 495 500

ATC TTC GCC CAG CGA CTG ATG CGA TAC TGG GCC AAC TTT GCC CGC ACA 1710
 Ile Phe Ala Gln Arg Leu Met Arg Tyr Trp Ala Asn Phe Ala Arg Thr
 505 510 515

GGG GAT CCC AAT GAG CCC CGA GAC CCC AAG GCC CCA CAA TGG CCC CCG 1758
 Gly Asp Pro Asn Glu Pro Arg Asp Pro Lys Ala Pro Gln Trp Pro Pro
 520 525 530

TAC ACG GCG GGG GCT CAG CAG TAC GTT AGT CTG GAC CTG CGG CCG CTG 1806
 Tyr Thr Ala Gly Ala Gln Gln Tyr Val Ser Leu Asp Leu Arg Pro Leu
 535 540 545

GAG GTG CGG CGG GGG CTG CGC GCC CAG GCC TGC GCC TTC TGG AAC CGC 1854
 Glu Val Arg Arg Gly Leu Arg Ala Gln Ala Cys Ala Phe Trp Asn Arg
 550 555 560 565

TTC CTC CCC AAA TTG CTC AGC GCC ACC GGT ATG CAG GGG CCA GCG GGC 1902
 Phe Leu Pro Lys Leu Leu Ser Ala Thr Gly Met Gln Gly Pro Ala Gly
 570 575 580

AGC GGC TGG GAG GAG GGG AGT GGG AGC CCG CCA GGT GTA ACC CCT CTC 1950
 Ser Gly Trp Glu Glu Gly Ser Gly Ser Pro Pro Gly Val Thr Pro Leu
 585 590 595

TTC TCC CCC TAGCCTCGGA GGCTCCCAGC ACCTGCCCAG GCTTCACCCA 1999
 Phe Ser Pro
 600

TGGGGAGGCT GCTCCGAGGC CCGGCCTCCC CCTGCCCCTC CTCCTCCTCC ACCAGCTTCT
 2059

CCTCCTCTTC CTCTCCCACC TCCGGCGGCT GTGAACACGG CCTCTTCCCC TACGGCCTAC
 2119

AGGGGCCCCT CCTCTAATGA GTGGTAGGAC CTGTGGGGAA GGGCCCCACT CAGGGATCTC
2179

AGACCTAGTG CTCCCTTCCT CCTCAAACCG AGAGACTCAC ACTGGACAGG GCAGGAGGAG
2239

GGGCCGTGCC TCCCACCCTT CTCAGGGACC CCCACGCCTT TGTTGTTTGA ATGGAAATGG
2299

AAAAGCCAGT ATTCTTTTAT AAAATTATCT TTTGGAACCT GAGCCTGACA TTGGGGGAAG
2359

TGGAGGCCCC GAAACGGGGT AGCACCCCCA TTGGGGCTAT AACGGTCAAC CATTTCTGTC
2419

TCTTC'TTTTT CCCCCAACCT CCCCCTCCTG TCCCCTCTGT TCCCGTCTTC CGGTCAATTCT
2479

TTTCTCCTCC TCTCTCCTT CTGCTGTCCT TCTCGGCCCC GCCTCTGCCC TCATCCTCCC
2539

TCTCGTCTTT CGCACATTCT CCTGATCCTC TTGCCACCGT CCCACGTGGT CGCCTGCATT
2599

TCTCCGTGCG TCCTCCCTGC ACTCATACCC CCCCTTCAAC CCGCCCCAAT GTCCGATCCC
2659

CGACCTTCCT CGTGCCGTCC TCCCCTCCCG CCTCGCTGGG CGCCCTGGCC GCAGACACGC
2719

TCGACACGCT CGACGAGGCG GAGCGCCAGT GGAAGGCCGA GTTCCACCGC TGGAGCTCCT
2779

ACATGGTGCA CTGGAAGAAC CAGTTCGACC ACTACAGCAA GCAGGATCGC TGCTCAGACC
2839

TGTGACCCCG GCGGGACCCC CATGTCCTCC GCTCCGCCCG GCCCCCTAGC TGTATATACT
2899

ATTTATTTCA GGGCTGGGCT ATAACACAGA CGAGCCCCAG ACTCTGCCCA TCCCCACCCC
2959

ACCCCGACGT CCCCCGGGGC TCCCGGTCTT CTGGCATGTC TTCAGGCTGA GCTCCTCCCC
3019

GCGTGCCTTC GCCCTCTGGC TGCAAATAAA CTGTTACAGG CCAAAAAAAAAA AAAAAAAAAA
3079

AAAAAAAAAA AAAAAA

3096

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 600 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Arg Pro Pro Gln Cys Leu Leu His Thr Pro Ser Leu Ala Ser Pro
1 5 10 15

Leu Leu Leu Leu Leu Leu Trp Leu Leu Gly Gly Gly Val Gly Ala Glu
20 25 30

Gly Arg Glu Asp Ala Glu Leu Leu Val Thr Val Arg Gly Gly Arg Leu
35 40 45

Arg Gly Ile Arg Leu Lys Thr Pro Gly Gly Pro Val Ser Ala Phe Leu
50 55 60

Gly Ile Pro Phe Ala Glu Pro Pro Met Gly Pro Arg Arg Phe Leu Pro
65 70 75 80

Pro Glu Pro Lys Gln Pro Trp Ser Gly Val Val Asp Ala Thr Thr Phe
85 90 95

Gln Ser Val Cys Tyr Gln Tyr Val Asp Thr Leu Tyr Pro Gly Phe Glu
100 105 110

Gly Thr Glu Met Trp Asn Pro Asn Arg Glu Leu Ser Glu Asp Cys Leu
115 120 125

Tyr Leu Asn Val Trp Thr Pro Tyr Pro Arg Pro Thr Ser Pro Thr Pro
130 135 140

Val Leu Val Trp Ile Tyr Gly Gly Gly Phe Tyr Ser Gly Ala Ser Ser
145 150 155 160

Leu Asp Val Tyr Asp Gly Arg Phe Leu Val Gln Ala Glu Arg Thr Val
165 170 175

Leu Val Ser Met Asn Tyr Arg Val Gly Ala Phe Gly Phe Leu Ala Leu
180 185 190

Pro Gly Ser Arg Glu Ala Pro Gly Asn Val Gly Leu Leu Asp Gln Arg
195 200 205

Leu Ala Leu Gln Trp Val Gln Glu Asn Val Ala Ala Phe Gly Gly Asp
210 215 220

Pro Thr Ser Val Thr Leu Phe Gly Glu Ser Ala Gly Ala Ala Ser Val
225 230 235 240

Gly Met His Leu Leu Ser Pro Pro Ser Arg Gly Leu Phe His Arg Ala

| | | |
|---|-----|-----|
| 245 | 250 | 255 |
| Val Leu Gln Ser Gly Ala Pro Asn Gly Pro Trp Ala Thr Val Gly Met | | |
| 260 | 265 | 270 |
| Gly Glu Ala Arg Arg Arg Ala Thr Gln Leu Ala His Leu Val Gly Cys | | |
| 275 | 280 | 285 |
| Pro Pro Gly Gly Thr Gly Gly Asn Asp Thr Glu Leu Val Ala Cys Leu | | |
| 290 | 295 | 300 |
| Arg Thr Arg Pro Ala Gln Val Leu Val Asn His Glu Trp His Val Leu | | |
| 305 | 310 | 315 |
| 320 | | |
| Pro Gln Glu Ser Val Phe Arg Phe Ser Phe Val Pro Val Val Asp Gly | | |
| 325 | 330 | 335 |
| Asp Phe Leu Ser Asp Thr Pro Glu Ala Leu Ile Asn Ala Gly Asp Phe | | |
| 340 | 345 | 350 |
| His Gly Leu Gln Val Leu Val Gly Val Val Lys Asp Glu Gly Ser Tyr | | |
| 355 | 360 | 365 |
| Phe Leu Val Tyr Gly Ala Pro Gly Phe Ser Lys Asp Asn Glu Ser Leu | | |
| 370 | 375 | 380 |
| Ile Ser Arg Ala Glu Phe Leu Ala Gly Val Arg Val Gly Val Pro Gln | | |
| 385 | 390 | 395 |
| 400 | | |
| Val Ser Asp Leu Ala Ala Glu Ala Val Val Leu His Tyr Thr Asp Trp | | |
| 405 | 410 | 415 |
| Leu His Pro Glu Asp Pro Ala Arg Leu Arg Glu Ala Leu Ser Asp Val | | |
| 420 | 425 | 430 |
| Val Gly Asp His Asn Val Val Cys Pro Val Ala Gln Leu Ala Gly Arg | | |
| 435 | 440 | 445 |
| Leu Ala Ala Gln Gly Ala Arg Val Tyr Ala Tyr Val Phe Glu His Arg | | |
| 450 | 455 | 460 |
| Ala Ser Thr Leu Ser Trp Pro Leu Trp Met Gly Val Pro His Gly Tyr | | |
| 465 | 470 | 475 |
| 480 | | |
| Glu Ile Glu Phe Ile Phe Gly Ile Pro Leu Asp Pro Ser Arg Asn Tyr | | |
| 485 | 490 | 495 |
| Thr Ala Glu Glu Lys Ile Phe Ala Gln Arg Leu Met Arg Tyr Trp Ala | | |
| 500 | 505 | 510 |
| Asn Phe Ala Arg Thr Gly Asp Pro Asn Glu Pro Arg Asp Pro Lys Ala | | |
| 515 | 520 | 525 |
| Pro Gln Trp Pro Pro Tyr Thr Ala Gly Ala Gln Gln Tyr Val Ser Leu | | |

530 535 540
 Asp Leu Arg Pro Leu Glu Val Arg Arg Gly Leu Arg Ala Gln Ala Cys
 545 550 555 560
 Ala Phe Trp Asn Arg Phe Leu Pro Lys Leu Leu Ser Ala Thr Gly Met
 565 570 575
 Gln Gly Pro Ala Gly Ser Gly Trp Glu Glu Gly Ser Gly Ser Pro Pro
 580 585 590
 Gly Val Thr Pro Leu Phe Ser Pro
 595 600

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3016 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 160..2010

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CCTCTCTCCC CTCATCTTTG CCAACCTGCC CCACCTCCTC TGCAGCTGAG CGATAACCCT
 60
 TGGGCCGACA GTGCCCTAAT CTCCTCCCTC CTGGCTTCTC GACCGACCCT TCACCCTTTC
 120
 CCTTTCTTTC TCCCAGCAGA CGCCGCCTGC CCTGCAGCC ATG AGG CCC CCG CAG 174
 Met Arg Pro Pro Gln
 605
 TGT CTG CTG CAC ACG CCT TCC CTG GCT TCC CCA CTC CTT CTC CTC CTC 222
 Cys Leu Leu His Thr Pro Ser Leu Ala Ser Pro Leu Leu Leu Leu Leu
 610 615 620
 CTC TGG CTC CTG GGT GGA GGA GTG GGG GCT GAG GGC CGG GAG GAT GCA 270
 Leu Trp Leu Leu Gly Gly Gly Val Gly Ala Glu Gly Arg Glu Asp Ala
 625 630 635
 GAG CTG CTG GTG ACG GTG CGT GGG GGC CGG CTG CGG GGC ATT CGC CTG 318
 Glu Leu Leu Val Thr Val Arg Gly Gly Arg Leu Arg Gly Ile Arg Leu
 640 645 650
 AAG ACC CCC GGG GGC CCT GTC TCT GCT TTC CTG GGC ATC CCC TTT GCG 366

| | |
|--|-----|
| Lys Thr Pro Gly Gly Pro Val Ser Ala Phe Leu Gly Ile Pro Phe Ala | |
| 655 660 665 | |
| GAG CCA CCC ATG GGA CCC CGT CGC TTT CTG CCA CCG GAG CCC AAG CAG | 414 |
| Glu Pro Pro Met Gly Pro Arg Arg Phe Leu Pro Pro Glu Pro Lys Gln | |
| 670 675 680 685 | |
| CCT TGG TCA GGG GTG GTA GAC GCT ACA ACC TTC CAG AGT GTC TGC TAC | 462 |
| Pro Trp Ser Gly Val Val Asp Ala Thr Thr Phe Gln Ser Val Cys Tyr | |
| 690 695 700 | |
| CAA TAT GTG GAC ACC CTA TAC CCA GGT TTT GAG GGC ACC GAG ATG TGG | 510 |
| Gln Tyr Val Asp Thr Leu Tyr Pro Gly Phe Glu Gly Thr Glu Met Trp | |
| 705 710 715 | |
| AAC CCC AAC CGT GAG CTG AGC GAG GAC TGC CTG TAC CTC AAC GTG TGG | 558 |
| Asn Pro Asn Arg Glu Leu Ser Glu Asp Cys Leu Tyr Leu Asn Val Trp | |
| 720 725 730 | |
| ACA CCA TAC CCC CGG CCT ACA TCC CCC ACC CCT GTC CTC GTC TGG ATC | 606 |
| Thr Pro Tyr Pro Arg Pro Thr Ser Pro Thr Pro Val Leu Val Trp Ile | |
| 735 740 745 | |
| TAT GGG GGT GGC TTC TAC AGT GGG GCC TCC TCC TTG GAC GTG TAC GAT | 654 |
| Tyr Gly Gly Gly Phe Tyr Ser Gly Ala Ser Ser Leu Asp Val Tyr Asp | |
| 750 755 760 765 | |
| GGC CGC TTC TTG GTA CAG GCC GAG AGG ACT GTG CTG GTG TCC ATG AAC | 702 |
| Gly Arg Phe Leu Val Gln Ala Glu Arg Thr Val Leu Val Ser Met Asn | |
| 770 775 780 | |
| TAC CGG GTG GGA GCC TTT GGC TTC CTG GCC CTG CCG GGG AGC CGA GAG | 750 |
| Tyr Arg Val Gly Ala Phe Gly Phe Leu Ala Leu Pro Gly Ser Arg Glu | |
| 785 790 795 | |
| GCC CCG GGC AAT GTG GGT CTC CTG GAT CAG AGG CTG GCC CTG CAG TGG | 798 |
| Ala Pro Gly Asn Val Gly Leu Leu Asp Gln Arg Leu Ala Leu Gln Trp | |
| 800 805 810 | |
| GTG CAG GAG AAC GTG GCA GCC TTC GGG GGT GAC CCG ACA TCA GTG ACG | 846 |
| Val Gln Glu Asn Val Ala Ala Phe Gly Gly Asp Pro Thr Ser Val Thr | |
| 815 820 825 | |
| CTG TTT GGG GAG AGC GCG GGA GCC GCC TCG GTG GGC ATG CAC CTG CTG | 894 |
| Leu Phe Gly Glu Ser Ala Gly Ala Ala Ser Val Gly Met His Leu Leu | |
| 830 835 840 845 | |
| TCC CCG CCC AGC CGG GGC CTG TTC CAC AGG GCC GTG CTG CAG AGC GGT | 942 |
| Ser Pro Pro Ser Arg Gly Leu Phe His Arg Ala Val Leu Gln Ser Gly | |
| 850 855 860 | |
| GCC CCC AAT GGA CCC TGG GCC ACG GTG GGC ATG GGA GAG GCC CGT CGC | 990 |
| Ala Pro Asn Gly Pro Trp Ala Thr Val Gly Met Gly Glu Ala Arg Arg | |
| 865 870 875 | |

| | |
|--|------|
| AGG GCC ACG CAG CTG GCC CAC CTT GTG GGC TGT CCT CCA GGC GGC ACT | 1038 |
| Arg Ala Thr Gln Leu Ala His Leu Val Gly Cys Pro Pro Gly Gly Thr | |
| 880 885 890 | |
| GGT GGG AAT GAC ACA GAG CTG GTA GCC TGC CTT CGG ACA CGA CCA GCG | 1086 |
| Gly Gly Asn Asp Thr Glu Leu Val Ala Cys Leu Arg Thr Arg Pro Ala | |
| 895 900 905 | |
| CAG GTC CTG GTG AAC CAC GAA TGG CAC GTG CTG CCT CAA GAA AGC GTC | 1134 |
| Gln Val Leu Val Asn His Glu Trp His Val Leu Pro Gln Glu Ser Val | |
| 910 915 920 925 | |
| TTC CGG TTC TCC TTC GTG CCT GTG GTA GAT GGA GAC TTC CTC AGT GAC | 1182 |
| Phe Arg Phe Ser Phe Val Pro Val Val Asp Gly Asp Phe Leu Ser Asp | |
| 930 935 940 | |
| ACC CCA GAG GCC CTC ATC AAC GCG GGA GAC TTC CAC GGC CTG CAG GTG | 1230 |
| Thr Pro Glu Ala Leu Ile Asn Ala Gly Asp Phe His Gly Leu Gln Val | |
| 945 950 955 | |
| CTG GTG GGT GTG GTG AAG GAT GAG GGC TCG TAT TTT CTG GTT TAC GGG | 1278 |
| Leu Val Gly Val Val Lys Asp Glu Gly Ser Tyr Phe Leu Val Tyr Gly | |
| 960 965 970 | |
| GCC CCA GGC TTC AGC AAA GAC AAC GAG TCT CTC ATC AGC CGG GCC GAG | 1326 |
| Ala Pro Gly Phe Ser Lys Asp Asn Glu Ser Leu Ile Ser Arg Ala Glu | |
| 975 980 985 | |
| TTC CTG GCC GGG GTG CGG GTC GGG GTT CCC CAG GTA AGT GAC CTG GCA | 1374 |
| Phe Leu Ala Gly Val Arg Val Gly Val Pro Gln Val Ser Asp Leu Ala | |
| 990 995 1000 1005 | |
| GCC GAG GCT GTG GTC CTG CAT TAC ACA GAC TGG CTG CAT CCC GAG GAC | 1422 |
| Ala Glu Ala Val Val Leu His Tyr Thr Asp Trp Leu His Pro Glu Asp | |
| 1010 1015 1020 | |
| CCG GCA CGC CTG AGG GAG GCC CTG AGC GAT GTG GTG GGC GAC CAC AAT | 1470 |
| Pro Ala Arg Leu Arg Glu Ala Leu Ser Asp Val Val Gly Asp His Asn | |
| 1025 1030 1035 | |
| GTC GTG TGC CCC GTG GCC CAG CTG GCT GGG CGA CTG GCT GCC CAG GGT | 1518 |
| Val Val Cys Pro Val Ala Gln Leu Ala Gly Arg Leu Ala Ala Gln Gly | |
| 1040 1045 1050 | |
| GCC CGG GTC TAC GCC TAC GTC TTT GAA CAC CGT GCT TCC ACG CTC TCC | 1566 |
| Ala Arg Val Tyr Ala Tyr Val Phe Glu His Arg Ala Ser Thr Leu Ser | |
| 1055 1060 1065 | |
| TGG CCC CTG TGG ATG GGG GTG CCC CAC GGC TAC GAG ATC GAG TTC ATC | 1614 |
| Trp Pro Leu Trp Met Gly Val Pro His Gly Tyr Glu Ile Glu Phe Ile | |
| 1070 1075 1080 1085 | |
| TTT GGG ATC CCC CTG GAC CCC TCT CGA AAC TAC ACG GCA GAG GAG AAA | 1662 |

Phe Gly Ile Pro Leu Asp Pro Ser Arg Asn Tyr Thr Ala Glu Glu Lys
1090 1095 1100

ATC TTC GCC CAG CGA CTG ATG CGA TAC TGG GCC AAC TTT GCC CGC ACA 1710
Ile Phe Ala Gln Arg Leu Met Arg Tyr Trp Ala Asn Phe Ala Arg Thr
1105 1110 1115

GGG GAT CCC AAT GAG CCC CGA GAC CCC AAG GCC CCA CAA TGG CCC CCG 1758
Gly Asp Pro Asn Glu Pro Arg Asp Pro Lys Ala Pro Gln Trp Pro Pro
1120 1125 1130

TAC ACG GCG GGG GCT CAG CAG TAC GTT AGT CTG GAC CTG CGG CCG CTG 1806
Tyr Thr Ala Gly Ala Gln Gln Tyr Val Ser Leu Asp Leu Arg Pro Leu
1135 1140 1145

GAG GTG CGG CGG GGG CTG CGC GCC CAG GCC TGC GCC TTC TGG AAC CGC 1854
Glu Val Arg Arg Gly Leu Arg Ala Gln Ala Cys Ala Phe Trp Asn Arg
1150 1155 1160 1165

TTC CTC CCC AAA TTG CTC AGC GCC ACC GCC TCG GAG GCT CCC AGC ACC 1902
Phe Leu Pro Lys Leu Leu Ser Ala Thr Ala Ser Glu Ala Pro Ser Thr
1170 1175 1180

TGC CCA GGC TTC ACC CAT GGG GAG GCT GCT CCG AGG CCC GGC CTC CCC 1950
Cys Pro Gly Phe Thr His Gly Glu Ala Ala Pro Arg Pro Gly Leu Pro
1185 1190 1195

CTG CCC CTC CTC CTC CTC CAC CAG CTT CTC CTC CTC TTC CTC TCC CAC 1998
Leu Pro Leu Leu Leu Leu His Gln Leu Leu Leu Phe Leu Ser His
1200 1205 1210

CTC CGG CGG CTG TGAACACGGC CTCTTCCCCT ACGGCCTACA GGGGCCCTC 2050
Leu Arg Arg Leu
1215

CTCTAATGAG TGGTAGGACC TGTGGGGAAG GGCCCCACTC AGGGATCTCA GACCTAGTGC
2110

TCCCTTCCTC CTCAAACCGA GAGACTCACA CTGGACAGGG CAGGAGGAGG GGCCGTGCCT
2170

CCCACCCTTC TCAGGGACCC CCACGCCTTT GTTGTTTGAA TGGAAATGGA AAAGCCAGTA
2230

TTCTTTTATA AAATTATCTT TTGGAACCTG AGCCTGACAT TGGGGGAAGT GGAGGCCCGG
2290

AAACGGGGTA GCACCCCCAT TGGGGCTATA ACGGTCAACC ATTTCTGTCT CTTCTTTTTC
2350

CCCCAACCTC CCCCTCCTGT CCCCTCTGTT CCCGTCTTCC GGTCATTCTT TTCTCCTCCT
2410

CTCTCCTTCC TGCTGTCCTT CTCGGCCCCG CCTCTGCCCT CATCCTCCCT CTCGTCTTTC
2470

GCACATTCTC CTGATCCTCT TGCCACCGTC CCACGTGGTC GCCTGCATTT CTCCGTGCGT
2530

CCTCCCTGCA CTCATACCCC CCCTTCAACC CGCCCAAATG TCCGATCCCC GACCTTCCTC
2590

GTGCCGTCCT CCCCTCCCGC CTCGCTGGGC GCCCTGGCCG CAGACACGCT CGACACGCTC
2650

GACGAGGCGG AGCGCCAGTG GAAGGCCGAG TTCCACCGCT GGAGCTCCTA CATGGTGCAC
2710

TGGAAGAACC AGTTCGACCA CTACAGCAAG CAGGATCGCT GCTCAGACCT GTGACCCCGG
2770

CGGGACCCCC ATGTCCTCCG CTCCGCCCGG CCCCCTAGCT GTATATACTA TTTATTTTCA
2830

GGCTGGGCTA TAACACAGAC GAGCCCCAGA CTCTGCCCAT CCCCACCCCA CCCCACGCTC
2890

CCCCGGGGCT CCCGGTCCTC TGGCATGTCT TCAGGCTGAG CTCCTCCCCG CGTGCCTTCG
2950

CCCTCTGGCT GCAAATAAAC TGTTACAGGC CAAAAAAAAA AAAAAAAAAA AAAAAAAAAA
3010

AAAAAA 3016

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 617 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Arg | Pro | Pro | Gln | Cys | Leu | Leu | His | Thr | Pro | Ser | Leu | Ala | Ser | Pro |
| 1 | | 5 | | 10 | | 15 | | | | | | | | | |
| Leu | Leu | Leu | Leu | Leu | Leu | Trp | Leu | Leu | Gly | Gly | Gly | Val | Gly | Ala | Glu |
| | 20 | | 25 | | 30 | | | | | | | | | | |
| Gly | Arg | Glu | Asp | Ala | Glu | Leu | Leu | Val | Thr | Val | Arg | Gly | Gly | Arg | Leu |
| | 35 | | 40 | | 45 | | | | | | | | | | |
| Arg | Gly | Ile | Arg | Leu | Lys | Thr | Pro | Gly | Gly | Pro | Val | Ser | Ala | Phe | Leu |

| | | |
|---|-----|-----|
| 50 | 55 | 60 |
| Gly Ile Pro Phe Ala Glu Pro Pro Met Gly Pro Arg Arg Phe Leu Pro | | |
| 65 | 70 | 80 |
| Pro Glu Pro Lys Gln Pro Trp Ser Gly Val Val Asp Ala Thr Thr Phe | | |
| 85 | 90 | 95 |
| Gln Ser Val Cys Tyr Gln Tyr Val Asp Thr Leu Tyr Pro Gly Phe Glu | | |
| 100 | 105 | 110 |
| Gly Thr Glu Met Trp Asn Pro Asn Arg Glu Leu Ser Glu Asp Cys Leu | | |
| 115 | 120 | 125 |
| Tyr Leu Asn Val Trp Thr Pro Tyr Pro Arg Pro Thr Ser Pro Thr Pro | | |
| 130 | 135 | 140 |
| Val Leu Val Trp Ile Tyr Gly Gly Gly Phe Tyr Ser Gly Ala Ser Ser | | |
| 145 | 150 | 160 |
| Leu Asp Val Tyr Asp Gly Arg Phe Leu Val Gln Ala Glu Arg Thr Val | | |
| 165 | 170 | 175 |
| Leu Val Ser Met Asn Tyr Arg Val Gly Ala Phe Gly Phe Leu Ala Leu | | |
| 180 | 185 | 190 |
| Pro Gly Ser Arg Glu Ala Pro Gly Asn Val Gly Leu Leu Asp Gln Arg | | |
| 195 | 200 | 205 |
| Leu Ala Leu Gln Trp Val Gln Glu Asn Val Ala Ala Phe Gly Gly Asp | | |
| 210 | 215 | 220 |
| Pro Thr Ser Val Thr Leu Phe Gly Glu Ser Ala Gly Ala Ala Ser Val | | |
| 225 | 230 | 240 |
| Gly Met His Leu Leu Ser Pro Pro Ser Arg Gly Leu Phe His Arg Ala | | |
| 245 | 250 | 255 |
| Val Leu Gln Ser Gly Ala Pro Asn Gly Pro Trp Ala Thr Val Gly Met | | |
| 260 | 265 | 270 |
| Gly Glu Ala Arg Arg Arg Ala Thr Gln Leu Ala His Leu Val Gly Cys | | |
| 275 | 280 | 285 |
| Pro Pro Gly Gly Thr Gly Gly Asn Asp Thr Glu Leu Val Ala Cys Leu | | |
| 290 | 295 | 300 |
| Arg Thr Arg Pro Ala Gln Val Leu Val Asn His Glu Trp His Val Leu | | |
| 305 | 310 | 320 |
| Pro Gln Glu Ser Val Phe Arg Phe Ser Phe Val Pro Val Val Asp Gly | | |
| 325 | 330 | 335 |
| Asp Phe Leu Ser Asp Thr Pro Glu Ala Leu Ile Asn Ala Gly Asp Phe | | |

| | | |
|---|-----|-----|
| 340 | 345 | 350 |
| His Gly Leu Gln Val Leu Val Gly Val Val Lys Asp Glu Gly Ser Tyr | | |
| 355 | 360 | 365 |
| Phe Leu Val Tyr Gly Ala Pro Gly Phe Ser Lys Asp Asn Glu Ser Leu | | |
| 370 | 375 | 380 |
| Ile Ser Arg Ala Glu Phe Leu Ala Gly Val Arg Val Gly Val Pro Gln | | |
| 385 | 390 | 400 |
| Val Ser Asp Leu Ala Ala Glu Ala Val Val Leu His Tyr Thr Asp Trp | | |
| 405 | 410 | 415 |
| Leu His Pro Glu Asp Pro Ala Arg Leu Arg Glu Ala Leu Ser Asp Val | | |
| 420 | 425 | 430 |
| Val Gly Asp His Asn Val Val Cys Pro Val Ala Gln Leu Ala Gly Arg | | |
| 435 | 440 | 445 |
| Leu Ala Ala Gln Gly Ala Arg Val Tyr Ala Tyr Val Phe Glu His Arg | | |
| 450 | 455 | 460 |
| Ala Ser Thr Leu Ser Trp Pro Leu Trp Met Gly Val Pro His Gly Tyr | | |
| 465 | 470 | 480 |
| Glu Ile Glu Phe Ile Phe Gly Ile Pro Leu Asp Pro Ser Arg Asn Tyr | | |
| 485 | 490 | 495 |
| Thr Ala Glu Glu Lys Ile Phe Ala Gln Arg Leu Met Arg Tyr Trp Ala | | |
| 500 | 505 | 510 |
| Asn Phe Ala Arg Thr Gly Asp Pro Asn Glu Pro Arg Asp Pro Lys Ala | | |
| 515 | 520 | 525 |
| Pro Gln Trp Pro Pro Tyr Thr Ala Gly Ala Gln Gln Tyr Val Ser Leu | | |
| 530 | 535 | 540 |
| Asp Leu Arg Pro Leu Glu Val Arg Arg Gly Leu Arg Ala Gln Ala Cys | | |
| 545 | 550 | 560 |
| Ala Phe Trp Asn Arg Phe Leu Pro Lys Leu Leu Ser Ala Thr Ala Ser | | |
| 565 | 570 | 575 |
| Glu Ala Pro Ser Thr Cys Pro Gly Phe Thr His Gly Glu Ala Ala Pro | | |
| 580 | 585 | 590 |
| Arg Pro Gly Leu Pro Leu Pro Leu Leu Leu His Gln Leu Leu Leu | | |
| 595 | 600 | 605 |
| Leu Phe Leu Ser His Leu Arg Arg Leu | | |
| 610 | 615 | |

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 45 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Leu Leu Ser Ala Thr Asp Thr Leu Asp Glu Ala Glu Arg Gln Trp Lys
1 5 10 15

Ala Glu Phe His Arg Trp Ser Ser Tyr Met Val His Trp Lys Asn Gln
20 25 30

Phe Asp His Tyr Ser Lys Gln Asp Arg Cys Ser Asp Leu
35 40 45

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 45 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Leu Leu Ser Ala Thr Asp Thr Leu Asp Glu Ala Glu Arg Gln Trp Lys
1 5 10 15

Ala Glu Phe His Arg Trp Ser Ser Tyr Met Val His Trp Lys Asn Gln
20 25 30

Phe Asp His Tyr Ser Lys Gln Glu Arg Cys Ser Asp Leu
35 40 45

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 48 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 45 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Leu Leu Ser Ala Thr Asp Thr Leu Asp Glu Ala Glu Arg Gln Trp Lys
1 5 10 15

Ala Glu Phe His Arg Trp Ser Ser Tyr Met Val His Trp Lys Asn Gln
20 25 30

Phe Asp His Tyr Ser Lys Gln Asp Arg Cys Ser Asp Leu
35 40 45

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 45 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Leu Leu Ser Ala Thr Asp Thr Leu Asp Glu Ala Glu Arg Gln Trp Lys
1 5 10 15

Ala Glu Phe His Arg Trp Ser Ser Tyr Met Val His Trp Lys Asn Gln
20 25 30

Phe Asp His Tyr Ser Lys Gln Glu Arg Cys Ser Asp Leu
35 40 45

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 48 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Leu Leu Ser Ala Thr Ala Ser Glu Ala Pro Ser Thr Cys Pro Gly Phe
1 5 10 15

Thr His Gly Glu Ala Ala Pro Arg Pro Gly Leu Pro Leu Pro Leu Leu
20 25 30

Leu Leu His Cys Leu Leu Leu Leu Phe Leu Ser His Leu Arg Arg Leu
35 40 45

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 47 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Leu Leu Ser Ala Thr Ala Thr Glu Val Pro Cys Thr Cys Pro Ser Pro
1 5 10 15

Ala His Gly Glu Ala Ala Pro Arg Pro Gly Pro Ala Leu Ser Leu Ser
20 25 30

Leu Leu Phe Phe Leu Phe Leu Leu His Ser Gly Leu Arg Trp Leu
35 40 45

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Leu Leu Ser Ala Thr Gly Met Gln Gly Pro Ala Gly Ser Gly Trp Glu
1 5 10 15

Glu Gly Ser Gly Ser Pro Pro Gly Val Thr Pro Leu Phe Ser Pro
 20 25 30

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 34 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Leu Leu Ser Ala Thr Gly Arg Arg Gly Val Gly Lys Gln Gly Met His
 1 5 10 15

Lys Ala Ala Arg Val Gly Arg Thr Gly Glu Arg Lys Gly Gly Lys His
 20 25 30

Arg Met

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CGGGTCTACG CCTACGTCTT TGAACACCGT GCTTC

35

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

CACAGGTCTG AGGAGCGATC CTGCTTGCTG

30

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GGTTAACTG GCGGGCTCC

19

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

ATGGGTGAAG CCTGGGCAGG TG

22

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

GCCCAGGCTT CACCCAT

17

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

AGACTGGGTA GATGATCAGA GACCTGAAAA CTACCG

36

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

GACAGGCCAG CTTGTGCTAT TGTTCAGAGT CTCAT

35

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

ACCGTCCACC TGAAGTGCTA CTGGGAGAAG

30

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

CGCTTACTAG GATCCAAGGC AAGCATGTAA

30

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

AGAAATGCAG GCGACCACGT G

21

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

CAGGTCCAGA CTAACGTACT GCTGAGCCCC CGCCG

35

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 374 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

ACACGCTCGA CGAGGCGGAG CGCCAGTGGG AGGCCGAGTT CCACCGCTGG AGTCCTACA
60

TGGTGCCTG GAAGAACCAG TTCGACCACT ACAGCAAGCA GGATCGCTGC TCAGACCTGT
120

GACCCCGGCG GGACCCCAT GTCCTCCGCT CCGCCCGGCC CCCTAGCTGT ATATACTATT
180

TATTCAGGG CTGGGCTATA ACACAGACGA GCCCCAGACT CTGCCCATCC CCACCCACC
240

CCGACGTCCC CCGGGGCTCC CGGTCCTCTG GCATGTCTTC AGGCTGAGCT CCTCCCCGCG
300

TGCCTTCGCC CTCTGGCTGC AAATAAACTG TTACAGGCCA AAAAAAAAAA AAAAAAAAAA
360

AAAAAAAAAA AAAA

374

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 40 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

Asp Thr Leu Asp Glu Ala Glu Arg Gln Trp Lys Ala Glu Phe His Arg
1 5 10 15

Trp Ser Ser Tyr Met Val His Trp Lys Asn Gln Phe Asp His Tyr Ser
20 25 30

Lys Gln Asp Arg Cys Ser Asp Leu
35 40

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1215 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 1..78

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

GGT ATG CAG GGG CCA GCG GGC AGC GGC TGG GAG GAG GGG AGT GGG AGC 48
Gly Met Gln Gly Pro Ala Gly Ser Gly Trp Glu Glu Gly Ser Gly Ser
620 625 630

CCG CCA GGT GTA ACC CCT CTC TTC TCC CCC TAGCCTCGGA GGCTCCCAGC 98
Pro Pro Gly Val Thr Pro Leu Phe Ser Pro
635 640

ACCTGCCCAG GCTTCACCCA TGGGGAGGCT GCTCCGAGGC CCGGCCTCCC CCTGCCCCCTC
158

CTCCTCCTCC ACCAGCTTCT CCTCCTCTTC CTCTCCCACC TCCGGCGGCT GTGAACACGG
218

CCTCTTCCCC TACGGCCTAC AGGGGCCCCCT CCTCTAATGA GTGGTAGGAC CTGTGGGGAA
278

GGGCCCCACT CAGGGATCTC AGACCTAGTG CTCCCTTCCT CCTCAAACCG AGAGACTCAC
338

ACTGGACAGG GCAGGAGGAG GGGCCGTGCC TCCCACCCTT CTCAGGGACC CCCACGCCTT
398

TGTTGTTTGA ATGGAAATGG AAAAGCCAGT ATTCTTTTAT AAAATTATCT TTTGGAACCT
458

GAGCCTGACA TTGGGGGAAG TGGAGGCCCG GAAACGGGGT AGCACCCCCA TTGGGGCTAT
518

AACGGTCAAC CATTCTGTC TCTTCTTTTT CCCCCAACCT CCCCCTCCTG TCCCCTCTGT
578

TCCCGTCTTC CGGTCATTCT TTTCTCCTCC TCTCTCCTTC CTGCTGTCCT TCTCGGCCCC 638

GCCTCTGCCC TCATCCTCCC TCTCGTCTTT CGCACATTCT CCTGATCCTC TTGCCACCGT
698

CCCACGTGGT CGCCTGCATT TCTCCGTGCG TCCTCCCTGC ACTCATACCC CCCCTTCAAC
758

CCGCCCAAAT GTCCGATCCC CGACCTTCCT CGTGCCGTCC TCCCCTCCCG CCTCGCTGGG
818

CGCCCTGGCC GCAGACACGC TCGACACGCT CGACGAGGCG GAGCGCCAGT GGAAGGCCGA
878

GTTCCACCGC TGGAGCTCCT ACATGGTGCA CTGGAAGAAC CAGTTCGACC ACTACAGCAA
938

GCAGGATCGC TGCTCAGACC TGTGACCCCG GCGGGACCCC CATGTCCTCC GCTCCGCCCC
998

GCCCCCTAGC TGTATATACT ATTTATTTCA GGGCTGGGCT ATAACACAGA CGAGCCCCAG
1058

ACTCTGCCCA TCCCCACCCC ACCCCGACGT CCCCCGGGGC TCCCGGTCCT CTGGCATGTC
1118

TTCAGGCTGA GCTCCTCCCC GCGTGCCTTC GCCCTCTGGC TGCAAATAAA CTGTTACAGG
1178

CCAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAA 1215

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

Gly Met Gln Gly Pro Ala Gly Ser Gly Trp Glu Glu Gly Ser Gly Ser
1 5 10 15

Pro Pro Gly Val Thr Pro Leu Phe Ser Pro
20 25